

# Automatic video mosaicing for surveillance using vanishing points

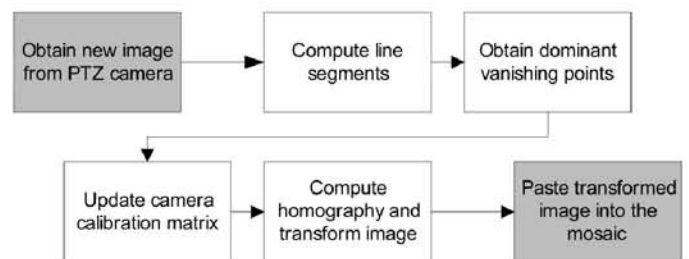
Marcos Nieto and Luis Salgado

*Moving cameras and projective geometry enable generation of real-time, user-controlled superimages of scenes.*

Video surveillance typically relies on various methods of 3D scene reconstruction, such as people counting, and object detection and tracking. Video mosaicing is a related technology that enables an expanded view of a scene by pasting together frames from the video stream (superimages) of a moving single camera in real time. This technique provides better resolution than can be achieved with a conventional video camera. Well-known related applications in other fields include composition of panoramic pictures, alignment of satellite images, and generation of synthetic 360°-angle views. Given two frames of a scene, a general approach consists of finding a number of points in both images that correspond to the same objects in the scene. This makes it possible to compute a ‘transform’ that describes the geometrical relation between the views. Known as a planar homography, this transform can be obtained using robust estimation techniques on the point correspondences.<sup>1</sup>

These strategies have the disadvantage that they use costly (i.e., time-consuming) image-processing algorithms for feature point extraction and matching, and iterative optimization algorithms to compute the homography. Here, we propose to take advantage of the moving nature of the cameras (which continuously change their rotation under the manipulation of an operator) to calibrate them and to use the resulting information to update, in real time, the video mosaic of the scene.

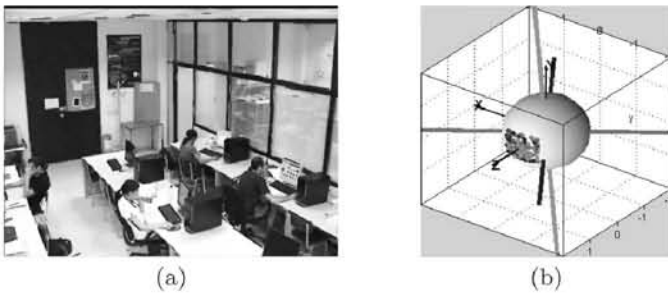
Although calibration is a well-known problem in the signal-processing literature,<sup>2</sup> there are no explicit solutions for automatic calibration for generic PTZ (pan, tilt, and zoom: the two rotation angles and the focal distance, which can be modified online by the user). This is a practical problem



*Figure 1. Block diagram of the image-processing chain. Each new image is pasted into the mosaic by updating the corresponding homography transform through the computation of vanishing points.*

for many typical scenarios in which PTZ cameras are components of interactive monitoring systems, as the interaction would invalidate any initial static calibration of the camera. Our approach takes advantage of the rotation-only nature of the PTZ environment to do the calibration (that is based on the computation of vanishing points). In this way, the homographies between views are completely defined by knowing the calibration matrix (given by the algorithm) and the rotation matrix (given by the known relative rotation of the camera). Figure 1 provides a block diagram of the proposed image-processing strategy.

Vanishing points refer to places in the images where parallel lines on the 3D world seem to converge. As projective entities, these points provide information about projected angles, and thus can be used to infer the calibration matrix associated with the camera. Specifically, it has been shown that three mutually orthogonal pairs of vanishing points (e.g., given by the three main directions of a building) contain enough information to recover the image of the absolute conic (a  $3 \times 3$  matrix containing five coefficients), from which a complete



**Figure 2.** Calculation of vanishing points. (a) Line segments are colored according to their vanishing point. (b) Sphere-centered view of the three dominant vanishing points.



**Figure 3.** Indoor example. The current view (a) is pasted into the mosaic (b) according to its estimated rotation and camera calibration.

camera calibration matrix can be retrieved using Cholesky decomposition.<sup>3</sup>

In our case, and to make computing less laborious, we do not assume that three vanishing points are present in the scene, rather we track a pair of dominant vanishing points (such as the vertical and one horizontal point) through a number of consecutive images produced by the movement of the PTZ camera. Three images with different rotations and the same pair of vanishing points gives us enough information to recover the image of the absolute conic.

The detection of vanishing points is carried out using a robust optimization method known as MLESAC (maximum likelihood estimation sample and consensus), which works on detected line segments of the images.<sup>4</sup> This method proceeds iteratively, searching for hypotheses of vanishing points that are verified by large amounts of line segments. The estimation of the vanishing points is provided by a nonlinear optimization algorithm that minimizes the angular error defined between the vanishing point and the direction of the line segments. Figure 2 shows the vanishing point detection for a typical indoor image.

For video mosaicing, the only required information (apart from the calibration of the camera) is the relative rotation matrix between two views. Typically, PTZ cameras work as Internet protocol servers that deliver specifications on request. Hence, we can use the pan and tilt angles to reconstruct the rotation matrix and the zoom to update the focal length parameter of the calibration matrix.

We have tested our vanishing point estimation algorithm against the most popular and recent methods in the literature and have shown that our approach is both highly accurate and significantly reduces the cost of computation. The result is an online application that works in real time (15 frames/second) and achieves calibration on structured scenarios for moving PTZ cameras and video mosaicing. Figure 3 shows an example of the method in action, generating an updated mosaic view (b) for each new incoming image (a). Our future efforts will combine computed online and offline reference calibrations using automatic techniques based on point correspondences.

## Author Information

**Marcos Nieto and Luis Salgado**

Grupo de Tratamiento de Imágenes

Universidad Politécnica de Madrid (UPM)

Madrid, Spain

Marcos Nieto received his telecommunication engineering degree from UPM in 2005. Since then, he has been a member of the Grupo de Tratamiento de Imágenes, where he is currently working toward his PhD in telecommunication engineering.

## References

1. R. Hartley and A. Zisserman, **Multiple View Geometry in Computer Vision**, Cambridge University Press, 2001.
2. R. Pflugfelder, **Self-Calibrating Cameras in Video Surveillance**, PhD thesis, University of Technology, Graz, 2008.
3. D. Liebowitz, **Camera Calibration and Reconstruction of Geometry**, PhD thesis, University of Oxford, 2001.
4. M. Nieto and L. Salgado, *Real-time robust estimation of vanishing points through nonlinear optimization*, **IS&T/SPIE Int. Conf. Real-Time Image Video Process.**, 2010.